Discovery of PPi-type Phosphoenolpyruvate Carboxykinase Genes in Eukaryotes and Bacteria

doi: 10.1074/jbc.M115.672907

Discovery of PPi-type phosphoenolpyruvate carboxykinase genes in eukaryotes and bacteria*

Yoko Chiba (千葉洋子)[1, 2], Ryoma Kamikawa (神川龍馬)[3], Kumiko Nakada-Tsukui (津久井久美子)[2], Yumiko Saito-Nakano (中野由美子)[2], Tomoyoshi Nozaki (野崎智義)[1, 2]

[1] Faculty of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan
[2] Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan
[3]Gradiate School of Environmental Studies, Graduate School of Human and Environmental Studies, Kyoto University, Yoshida Nihonmatsu cho, Kyoto, Kyoto 606-8501, Japan

*Running title: *Identification of PPi-type phosphoenolpyruvate carboxykinase*

To whom correspondence should be addressed: Tomoyoshi Nozaki, Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Tel.: +81 3 4580 2690; fax: +81 3 5285 1173. nozaki@nih.go.jp

Keywords: carbon fixation, carbohydrate metabolism, central metabolism, enzyme, novel enzyme, pyrophosphate, parasite metabolism, phosphoenolpyruvate carboxykinase

Background: Inorganic pyrophosphate-type phosphoenolpyruvate carboxylase (PPi-PEPCK) was unidentified.
Results: A conserved hypothetical protein was annotated as PPi-PEPCK.
Conclusion: PPi-PEPCK arose independently from the functional homologs, ATP/GTP-PEPCKs or PEP carboxylase.
Significance: Identification of PPi-PEPCK reveals the wide distribution of this enzyme and accelerates understanding the diversity of the central metabolism.

ABSTRACT
Phosphoenolpyruvate carboxykinase (PEPCK) is one of the pivotal enzymes, which regulates the carbon flow of the central metabolism by fixing CO_2 to phosphoenolpyruvate (PEP) to produce oxaloacetate or vice versa. While ATP- and GTP-type PEPCKs have been well studied and their protein identities are established, inorganic pyrophosphate (PPi)-type PEPCK (PPi-PEPCK) is poorly characterized. Despite extensive enzymological studies, its protein identity and encoding gene remain unknown. In this study, PPi-PEPCK has been identified for the first time from an eukaryotic human parasite, *Entamoeba histolytica,* by conventional purification and mass spectrometric identification of the native enzyme, followed by demonstration of its enzymatic activity. A homolog of the amebic PPi-PEPCK from an anaerobic bacterium *Propionibacterium freudenreichii* subsp. *shermanii* also exhibited PPi-PEPCK activity. The primary structure of PPi-PEPCK has no similarity to the functional homologs ATP/GTP-PEPCKs or PEP carboxylase, strongly suggesting that PPi-PEPCK arose independently from the other functional homologues and highly likely has unique catalytic sites. PPi-PEPCK homologs were found in a variety of bacteria and some eukaryotes, but not in archaea. The molecular identification of this long-forgotten enzyme tells us the diversity and functional redundancy of enzymes involved in the central metabolism and would help us to understand the central metabolism more deeply.

Inorganic pyrophosphate (PPi) is composed of two molecules of phosphate (Pi) linked by a phosphoanhydride bond. PPi has been proposed as an evolutionary precursor of ATP and GTP (1) since this structurally simple compound can be formed spontaneously and the hydrolysis of PPi produces high energy (2,3). If ancestral organisms

could utilize only PPi or polyphosphates as an energy and/or phosphate donor for enzymatic reactions, there should have been an event in which new enzymes utilizing ATP/GTP arose. This further raises the question of how organisms change the major substrate; by accumulation of mutation on the enzyme to change the substrate specificity from PPi to nucleotide triphosphate or the substitution of PPi-utilizing enzymes by an independently emerged functional homolog that utilize nucleotide triphosphates. To gain insight into the evolutionary transition from PPi to nucleotide triphosphate, evolutionary relationship between the extant PPi- and ATP/GTP- utilizing enzymes has been examined (4-11).

In glycolytic/gluconeogenic pathways and the closely related reactions in current organisms, three enzymatic reactions that can utilize PPi as the substrate have been reported. The first one is PPi-dependent phosphofructokinase (PFK) (EC 2.7.1.90) reaction. PPi-PFK catalyzes reversible reaction while ATP-dependent PFK (EC 2.7.1.11) catalyze irreversible reaction.

PPi + D-fructose 6-phosphate <=> Pi + D-fructose 1,6-bisphosphate
ATP + D-fructose 6-phosphate => ADP + D-fructose 1,6-bisphosphate

PPi-PFKs from an eukaryotic human parasite *Entamoeba histolytic*a (12) and an anaerobic bacterium *Propionibacterium freudenreichii* subsp. *shermanii* (13) were proposed to work for fructose 1,6-bisphosphatase synthesis like ATP-PFKs. In higher plants, it has been shown that PPi-PFK works in the opposite direction at least during internode developmental stages (14). Therefore, PPi-PFK can catalyze both PPi-utilizing and PPi-producing reactions not only *in vivo* but also *in vitro*. ATP- and PPi-PFK share a common origin, but the evolutionary history is highly complex. Change of the phosphate donors in PFKs occurred more than once, as suggested by phylogenetic analyses of ATP- and PPi-PFKs (6,9,10,15). The complex evolution of substrate utility in PFK suggests, at least in PFK, the transition of PPi-utilizing and nucleotide triphosphate-utilizing ability can have occurred relatively easily. Indeed, just a few-amino acid mutation can change the substrate preference from PPi to ATP based on the ratio of k_{cat}/K_m (8).

The second reaction utilizing PPi is catalyzed by pyruvate phosphate dikinase (PPDK; EC 2.7.9.1):

PPi + AMP + PEP <=> Pi + ATP + pyruvate

PEP–pyruvate conversion is also catalyzed by PPi-independent enzymes, PEP synthase (EC 2.7.9.2: upper formula of the below) and pyruvate kinase (EC 2.7.1.40: lower formula of the below):

Pi + AMP + PEP <=> ATP + pyruvate
ADP + PEP => ATP + pyruvate

PPDK and PEP synthase share conserved domains while overall amino acid sequence similarity is not high (16). On the other hand, neither PPDK nor PEP synthase shows substantial similarity to pyruvate kinase at the primary structure level. This case is apparently distinguishable from that of PFK because transition of substrate utility for the PEP-pyruvate conversion has not been happened by simple amino acid substitutions as seen in PFK.

The third mechanism is PEP-oxaloacetate interconverting reaction catalyzed by phosphoenolpyruvate carboxykinase (PEPCK). Because PEP is a key intermediate in a variety of metabolic processes in all living organisms (17,18), PEPCK works as a major crossroad that connects glycolysis/gluconeogenesis and organic acids metabolisms like TCA cycle and fumarate fermentation. According to the phosphate donor to oxaloacetate, PEPCK can be divided into three types: GTP- (EC 4.1.1.32), ATP- (EC 4.1.1.49), and PPi- (EC 4.1.1.38). PEPCK reactions are fundamentally reversible; however, ATP- and GTP-PEPCKs prefer a PEP-producing reaction whereas PPi-PEPCK prefers an oxaloacetate-producing reaction at least *in vitro* (19). PEP carboxylase (EC 4.1.1.31) also catalyzes PEP-oxaloacetate interconversion, but this reaction is irreversible and requires HCO_3^- instead of CO_2 (18,20).

GTP + oxaloacetate <=> GDP + PEP + CO_2
ATP + oxaloacetate <=> ADP + PEP + CO_2
PPi + oxaloacetate <=> Pi + PEP + CO_2
Pi + oxaloacetate <= PEP + HCO_3^-

ATP- and GTP-PEPCKs and PEP carboxylase have been well characterized. ATP-PEPCKs are mainly present in bacteria, yeast and plants, while GTP-PEPCKs are mostly present in higher eukaryotes, archaea and some bacteria (21). PEP carboxylase exists in various bacteria, and limited archaea and eukaryotes. Although there is no discernible similarity in the overall structure of these proteins, the residues implicated for the binding to metal, oxaloacetate/PEP, and nucleotides are conserved between ATP- and GTP-PEPCKs (22,23), while conservation of these residues has not been reported in the case of PEP carboxylases. In contrast, the information on PPi-PEPCK, which is also called as PEP carboxytransphosphorylase, is limited. Although PPi-PEPCK was previously purified from *P. freudenreichii* subsp. *shermanii* (19,24-26) and *E. histolytica* (27) and extensively characterized enzymologically, the gene encoding PPi-PEPCK remaines unknown since the first description in 1961 (24). Both the distribution of PPi-PEPCK among three domains of life and the evolutionary relationship between other PEP–oxaloacetate interconverting enzymes were totally obscure.

In the present study, we identified the protein for PPi-PEPCK activity and its encoding gene from *E. histolytica* and *P. freudenreichii*. The identification of PPi-PEPCK finally enabled us to compare the amino-acid sequences of all the three types of PEPCKs as well as PEP carboxylase. It also allowed us to estimate the distribution of PPi-PEPCK in the tree of life. Lastly, we argued the complex evolutionary history of proteins utilizing different substrates for the same catalytic reaction.

EXPERIMENTAL PROCEDURES

Microorganisms and cultivation– Trophozoites of *E. histolytica* clonal strain HM-1: IMSS Cl6 was maintained axenically in Diamond's BI-S-33 medium (28) at 35.5°C as described previously. Cells were grown to the late-logarithmic phase (2-3 days after inoculation), harvested by the addition of ice-cold PBS buffer to culture flasks, after discarding the medium, and followed by centrifugation at 300 × g for 5 min at 4°C. For protein purification, the harvested cell pellets were stored at -80 °C until use.

Enzyme assays– PPi and oxaloacetate-producing direct PEPCK activity was assayed by measuring the oxidation of NADH by detecting the decrease of absorbance at 340 nm as described previously (27). Reaction mixture contained 10 mM potassium phosphate buffer (pH 6.5), 0.4 mM PEP, 0.1 mM $CoCl_2$, 20 mM $KHCO_3$, 10 mM $MgCl_2$, 2.0 U mL^{-1} malate dehydrogenase (from rabbit muscle; Sigma), 0.25 mM NADH, and an enzyme solution. Reaction was started by adding PEP, and the NADH oxidizing activity without the addition of PEP was subtracted as the base line. To get the absolute active value, 400 μL of the reaction mixture was added in 1.0 cm path-length cuvettes and absorbance was monitored by a spectrometer. To get the relative activity during the purification, 100 μL of the mixture was put in a 96 well plate, incubated at 37˚C, and 340 nm was monitored by a micro-plate reader. One unit of activity was defined as the amount of enzyme oxidizing 1 μmol of NADH min^{-1}.

Purification of native E. histolytica *PEPCK*– EhPEPCK was purified from 5 g of wet cells as follows. The cells were suspended in 15 mL of 20 mM Tris-HCl (pH 8.0) containing 0.1 mM $CoCl_2$ (buffer A) and 0.5 mg mL^{-1} E-64, disrupted by freezing and thawing, and cell debris was then removed by centrifugation at 100,000 × *g* for 1 h. After the supernatant was diluted to 30 mL with buffer A, ammonium sulfate was added to give 30% saturation. The samples were then applied to a Butyl-Toyopearl column (22 mm × 15 cm; Tosoh, Tokyo, Japan) equilibrated with buffer A supplemented with ammonium sulfate at 30% saturation. All chromatography steps were performed using an ÄKTA purifier system (GE Healthcare, NJ, USA) at room temperature. Proteins were eluted with a gradient of ammonium sulfate from 30% to 0% at a flow rate of 4 mL

min^{-1}. The active fractions were desalted using PD10 column (GE Healthcare) and then applied to a MonoQ HR 5/5 column (bed volume, 1 mL; GE Healthcare) equilibrated with buffer A. Proteins were eluted with a gradient of NaCl from 0 to 500 mM for 20 column volume at a flow rate of 1.0 mL min^{-1}. The active fractions were pooled and loaded onto a Superdex 200 (10/300) column (GE Healthcare). Proteins were eluted with buffer A supplemented with 150 mM NaCl.

Identification of EhPEPCK by LC-MS/MS analysis— ~130 kDa bands were excised from silver stained gel and subjected to the LC-MS/MS analysis, at W. M. Keck Biomedical Mass Spectrometry Laboratory, University of Virginia, USA. The analysis was performed almost the same as before (29); only the difference was that 5 μL of the extract was injected to a reversed-phase capillary column and peptides were eluted over 0.3 hours.

To estimate the relative ratio of EhPEPCK1, 2 and 3 in the analyzed sample, the peak area for the selected ion chromatogram of the monoisotopic mass of the most abundant charge state (+/- 0.02Da) for each peptide was quantified. Three peptides conserved to all the three proteins, specific to EhPEPCK1, and conserved only in EhPEPCK2 and 3, respectively and one peptide unique to EhPEPCK2 or 3, respectively were selected for the analysis because these peptides were detected from all the three samples; immunoprecipitated samples (see below) using EhPEPCK1-HA, EhPEPCK2-HA, and purified PPi-PEPCK from the wild-type ameba lysate. Sum of the peak area value of the three peptides conserved to all was used for normalization.

Construction of plasmids– For expression in *Escherichia coli,* the protein coding sequence of *EhPEPCK1* (XP_654765.1) and *2* (XP_650862.1) genes were PCR-amplified from *E. histolytica* cDNA using a pair of following forward and reverse primers: 5’-TCGAAGGTAGGCATATGTTTAATCAAGA AAAAGGTACC-3’ (PEPCK_1_pCold_F) and 5’-ATTCGGATCCCTCGATTAATGTTTCATGC ATTTGTATG-3’ (PEPCK_1_2_pCold_R) for PEPCK1; 5’-TCGAAGGTAGGCATATGTTTAATCAAGA ACAAGGTA-3’ (PEPCK_2_pCold_F), and PEPCK_1_2_pCold_R for PEPCK2. For expression in *E. histolytica* trophozoites with hemagglutinin (HA) tag at the C-termini, the protein coding sequence of *EhPEPCK1* and *2* genes were PCR-amplified using the following primers: 5’-ACACATTAACAGATCATGTTTAATCAAG AAAAAGGTAC-3’ (PEPCK_1_pEhExHA_F) and 5’-ATGGATACATAGAATGTTTCATGCATTTG TATG-3’ (PEPCK_1_2_pEhExHA_R); and 5’-ACACATTAACAGATCATGTTTAATCAAGAA CAAGGTA -3’ (PEPCK_2_pEhExHA_F), and PEPCK_1_2_pEhExHA_R, respectively. pColdI (Takara, Japan) and pEhExHA (30) plasmids were used for the expression in *E. coli* and *E. histolytica*, respectively. The amplified fragments were inserted into the plasmids cut with NdeI and XhoI (pColdI) or BglII (pEhExHA) using In-Fusion HD Cloning system (Takara, Japan).

PPi-PEPCK gene from *P. freudenreichii* subsp. *freudenreichi*i (PfPEPCK; LC062511) was PCR-amplified from the genomic DNA purchased from German Collection of Microorganisms and Cell Cultures (DSM 20271; DSMZ, Germany) using the following primers: 5’-TCGAAGGTAGGCATAATGTCCGTAGTCG AACGC-3’ (Pf_PEPCK_pCold_F) and 5’-ATTCGGATCCCTCGATCAGACGAACCTG GGCTG -3’ (Pf_PEPCK_pCold_R); and cloned into pColdI as described above.

Overexpression and purification of recombinant PPi-PEPCKs–The plasmids for bacterial expression were introduced into *E. coli* BL21 CodonPlus (DE3)-RIL (EhPEPCK1 and 2; Agilent) of and *E. coli* BL21 Star (DE3) (PfPEPCK; Life technologies). The hosts transformed with the expression plasmids were inoculated into 400 mL or Luria-Bertani medium in 1 L conical flask containing 50 μg mL^{-1} ampicillin and 34 μg mL^{-1} chloramphenicol, if necessary. After cultivating the cells aerobically at 37 °C until the OD_{600} reached approximately 0.5, protein expressions were induced by cooling the culture on ice for 30 min and adding 0.1 mM

isopropyl thio-β-d-galactopyranoside to the medium, followed by over-night cultivation at 15 °C. Harvested cells (~8 g of wet cells from 1.2 L of culture) were disrupted by adding BugBuster (~5 mL g^{-1} of wet cells; Merck) and cell debris was removed by centrifugation. Imidazole and NaCl concentrations in the supernatant were adjusted to 10 and 300 mM, respectively. The supernatant was then applied to an open column packed with Ni-NTA agarose (1.5 mL of bed volume; Qiagen). After washing the column with 16 bed volume of 50 mM Tris-HCl, 300 mM NaCl, and 20 mM imidazole-HCl (pH 8.0), the His-tagged protein was eluted with 4 bed volume of the same buffer containing 250 mM imidazole-HCl. The eluted proteins were further purified using MonoQ column as described above.

Antibodies– To make anti-EhPEPCK antisera, full-length EhPEPCK1 with His-tag at N terminus was expressed in *E. coli*, purified using Ni-NTA and MonoQ columns as described above, and used as the antigen. The anti-EhPEPCK1 antisera from rabbit were commercially raised by Operon Biotechnology (Tokyo, Japan). The specificity of these antibodies was confirmed using *E. histolytica* lysates by western blotting (WB) analysis (data not shown).

Ameba transformation– The plasmids generated as described above were introduced into *E. histolytica* trophozoites by lipofection, as previously described (31) with minor modifications. Approximately 40 µl of the transfection medium containing 3-5 µg of one of the plasmids was mixed with 10 µl of PLUS™ reagent (Life Technologies) and kept at room temperature for 15 min. This mixture was combined with 20 µl of Lipofectamine™ transfection reagent mixed with 30 µL of transfection medium, kept at room temperature for 15 min, diluted with 400 µl of transfection medium, and added to the seeded trophozoites attached to 12 wells-plate after removing medium. The plate was then incubated at 35.5 °C for 5 h, and cells were transfer to a tube containing 5.5 mL of BIS medium. About 24 h after transfection, BIS medium was changed into fresh one with 1 µg/ mL of G418, and gradually increased for ~ 2 weeks until the G418 concentration reached 10 µg/ mL.

Protein assay– Protein concentrations were measured using a Bio-Rad protein assay DC dye. Bovine gamma globulin was used as a standard.

WB– Protein samples were subjected to SDS-PAGE after boiled at 95˚C for 3 min with SDS-PAGE loading buffer. Proteins in the gel were transferred to nitrocellulose membrane and then blocked with 5% (w/v) skim milk in Tris buffered saline with Tween20 (TBST) for overnight at 4 ˚C. Proteins on the membrane were reacted with antibodies in TBST for 1 h at room temperature. Primary antibodies were used at a 1:5000 dilution for anti-EhPEPCK rabbit antibody, a 1:1000 dilution for anti-cysteine synthase 1 (32) and anti-Cpn60 (33) rabbit antibodies, and anti-HA mouse monoclonal antibody (clone 11MO, Covance, Princeton, NJ, USA). The blots were visualized using alkaline phosphatase-conjugated anti-rabbit or mouse IgG antibody (1:2000 dilution; Cell Signaling Technology, USA) with AP Color Reagent (Bio Rad, USA) according to the manufacturer's protocol.

Immunoprecipitation– To examine the interaction between EhPEPCKs, 0.08 g of wet cells expressing PEPCK1-HA, PEPCK2-HA or only HA were lysed with 1 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1 mM $CoCl_2$, 0.2% TritonX-100, 0.5 mg mL^{-1} of E-64) for 10 min on ice. After centrifugation at 20,000 × g for 10 min, the supernatants were pre-cleaned with 150 µL of Protein G Sepharose (GE Healthcare, Waukesha, WI, USA) for 1h. After removing the beads by centrifugation, the precleared supernatant samples were mixed and incubated with 50 µL of anti-HA monoclonal antibody-conjugated agarose (Sigma-Aldrich, St. Louis, MO, USA) for 4 h. The agarose beads were washed with 400 µL of lysis buffer without E-64 for 5 times (9,300 × g for 1 min each) to remove the unbound proteins, and the bound proteins were eluted by incubating with 200 µL of lysis buffer containing 0.2 mg mL^{-1} of HA peptide (Sigma-Aldrich) for over-night. The eluted samples were collected by centrifugation at 800 × g for 3 min. All the centrifugations and

incubations were done at 4˚C otherwise mentioned.

Phylogenic analysis– Homologues of PPi-PEPCK were searched by the blastP analysis against non-redundant protein sequence database using EhPEPCK1 as a query (5th May 2015). In addition, we also retrieved the homologues of *Mastigamoeba balamuthi* and *Spironucleus barkhanus* by the tblastN analysis against the whole genome shotgun reads and the EST database, respectively. Other eukaryotic counterparts were retrieved by the tblastN analysis against the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) database (34). No more than two sequences with 97-100% amino acid sequence similarity were included, except for *Entamoeba* spp. and *Giardia intestinalis* (see discussion for the details). Sequences were aligned by MAFFT (35), and ambiguously aligned sites were excluded, resulting in the dataset comprising 158 taxa 749 positions. The dataset was subjected to the ML method with the LG + Γ + F model. The ML tree was heuristically searched from ten distinct maximum parsimony (MP) trees. In ML bootstrap analyses (100 replicates), a heuristic tree search was performed from a single MP tree per replicate. The ML phylogenetic analyses described above were conducted by RAxML ver. 7.2.8 (36).

RESULTS

Identification of PPi-PEPCK from E. histolytica– *In silico* search of potential PEPCK homologs in the *E. histolytica* genome (AmoebaDB; http://amoebadb.org/amoeba/) using canonical ATP- and GTP-PEPCK from *Dictyostelium discoideum* (XP_645490.1 and XP_645396.1, respectively) failed to detect any candidate of PEPCK (data not shown) although PPi-PEPCK activity was detected from *E. histolytica* trophozoite lysates (0.165 U mg of protein^{-1}) as previously reported by Reeves (27). PPi-PEPCK activity was sequentially separated using hydrophobic (Butyl-Toyopeal) and anion-exchange (MonoQ) columns (Table 1) and a MonoQ fraction containing peak PPi-PEPCK activity gave one major protein band at ~130 kDa on SDS-PAGE analysis (Fig. 1A). PPi-PEPCK activity detected in this MonoQ fraction indicated 206-fold purification with a specific activity of 34.0 U mg^{-1}, which well agrees with the previous results of PPi-PEPCK from *P. freudenreichii* that showed 12.8 U mg^{-1} of activity after 99-fold purification (25). The activity level was also similar to ATP-PEPCK from *Escherichia coli* (24 U mg^{-1}) (37). Phosphate, PEP, and MDH dependence of the NADH oxidizing activity in the MonoQ fraction was also confirmed. NADH oxidizing activity was not detected when phosphate was substituted to ADP, indicating that the fraction did not contain ATP-PEPCK activity.

The fraction containing PEPCK activity was further subjected to size-exclusion chromatography using Superdex 200. The PPi-PEPCK activity was detected in the fractions corresponding to two major peaks around ~143 and ~440 kDa (Fig. 2). The integrated areas of the absorbance at 280 nm corresponding to the 440-kDa and 143-kDa proteins (the first and second peaks, respectively) and their enzymatic activity levels are almost comparable, suggesting that active PPi-PEPCK is present as both monomer and trimer in *E. histolytica* (Fig. 2A). PPi-PEPCK from *P. freudenreichii* is previously reported to exist as monomer, dimer and tetramer with activity (26). Formation of several quaternary structures may be a character of PPi-PEPCK.

Since both the fractions seemed to contain an identical ~130-kD protein band on SDS-PAGE, the protein band from the fractions corresponding to the first peak (~440 kDa) was excised and subjected to ToFMS/MS analysis. Peptides detected from the band correspond to a “cluster of hypothetical proteins”, which consist of three homologous proteins: EHI_166920 (XP_654765.1), EHI_030750 (XP_650862.1), and EHI_198620 (XP_655201.1) (Table 2). The calculated molecular mass of these proteins (131 kDa) well agreed with that expected from the SDS-PAGE result, and the detected peptides covered 63-67% of the three proteins. Hereinafter these proteins are designated as EhPEPCK1, 2, and 3, respectively. EhPEPCK1 and 2 or 3 showed 91% amino acid identity, while EhPEPCK2 and 3 showed 98% identity. Amino acid variations

between EhPEPCK1 and EhPEPCK2/3 were mainly found in the first 60 residues at the amino-terminal region, showing only 45% identity between EhPEPCK1 and EhPEPCK2/3. Neither of these proteins show any similarity to known proteins nor domains as predicted by Pfam (http://pfam.xfam.org/), indicating that the purified proteins were novel proteins.

Confirmation of PPi-PEPCK activity of the amebic enzymes expressed in E. coli– EhPEPCK1, EhPEPCK2, and PfPEPCK were expressed as soluble recombinant proteins with the histidine tag at the amino terminus in *E. coli*. A Coomassie Brilliant Blue-stained SDS-PAGE gel of the proteins purified using Ni-NTA and MonoQ chromatography showed that the full length proteins of the expected size were purified to homogeneity (Fig. 1B). PPi-PEPCK activity of the purified EhPEPCK1, EhPEPCK2, and PfPEPCK were 8.9, 3.9, and 23.0 U mg^{-1}, respectively, confirming that the identified genes indeed encode proteins with PPi-PEPCK activity.

Expression of EhPEPCK1 and 2 with HA-tag in E. histolytica– Amebic transformants expressing either EhPEPCK1 or EhPEPCK2 with the HA tag at the carboxyl terminus (EhPEPCK1-HA or EhPEPCK2-HA) using the episomal plasmid were established. Expression of the epitope-tagged proteins were confirmed by WB using anti-HA antibody (Fig. 3A). The molecular mass estimated from the band corresponding to EhPEPCK1-HA was slightly larger than that of EhPEPCK2-HA. We hypothesize that this subtle difference in the mobility between EhPEPCK1 and 2 is attributable to the difference in the isoelectric point between the two proteins. The isoelectric point values of EhPEPCK1 and 2 are calculated as 6.15 and 6.06, respectively. At this stage, we do not know if the slight difference in the isoelectric point and the molecular size between EhPEPCK1 and 2 is physiologically important.

Heteromeric configuration of EhPEPCKs– To investigate whether EhPEPCKs form homo- or hetero-trimer, EhPEPCK1 and 2 were immuno-precipitated using lysates from transformants expressing corresponding EhPEPCK with the HA tag using anti-HA antibody. Both of the immuno-precipitated samples contained peptides from all the three EhPEPCKs (Table 3). When proteins were immuno-precipitated using EhPEPCK1-HA, EhPEPCK2/3 specific peptides were also detected while the relative ratio decreased about 1.7 folds compared to that of trimetric EhPEPCK purified from the wild-type *E. histolytica*. Similarly, the immuno-precipitated sample using EhPEPCK2-HA also contained peptides unique to EhPEPCK1 while the relative ratio decreased 2.0 folds. These data suggest that some proportion of the trimer consists as the mixture of EhPEPCK1, 2, and 3. The detailed composition of the trimeric complex should be analyzed in future.

Detection of PPi-PEPCKs by WB using anti-EhPEPCK antiserum– The anti-EhPEPCK antiserum reacted with recombinant PEPCK1 and 2 with comparable efficiency (Fig. 3B). Major bands of the predicted molecular weight (~130 kDa) were observed in *E. histolytica*'s lysate (Fig. 3C), confirming the specificity of this anti-EhPEPCK antiserum. We estimated that EhPEPCKs account for approximately 1% of the total protein of the amebic lysate by comparing the band intensity of serially diluted recombinant PEPCKs and amebic lysates in WB analysis (Fig. 3B). This estimation also agrees well with the result that about 200-folds purification of *E. histolytica* lysate gives highly purified PPi-PEPCK (Table 1 and Fig. 1A).

Localization of PPi-PEPCKs in E. histolytica *trophozoites*– *E. histolytica* trophozoites were mechanically disrupted in sucrose-MOPS buffer and centrifuged to separate the cytosol and organelle/membrane fractions. The cytosolic marker, cysteine synthase 1 and the organelle marker, Cpn60 were demonstrated in the corresponding fractions (Fig. 4). When these fractions were reacted with anti-EhPEPCK antiserumn, clear bands were detected from the supernatant fraction. Therefore, major parts of both EhPEPCK1 and 2/3 localize in the cytosol (Fig. 4), while two isoforms of GTP-PEPCK exist in the cytosol and mitochondria in case of animals (38).

In silico analysis to find organelle targeting

signals (TargetP 1.1, http://www.cbs.dtu.dk/services/TargetP/; SignalP 4.1, http://www.cbs.dtu.dk/services/SignalP/), or transmembrane regions (TMHMM Server v. 2.0, http://www.cbs.dtu.dk/services/TMHMM/) did not detect any targeting signals and thus were consistent with the result from the wet experiment described above.

Distribution of PPi-PEPCK in the tree of life– Homologs of EhPEPCKs were found in limited lineages of eukaryotes and bacteria, but not in archaea (Table 4 and 5). PPi-PEPCK genes were found in all the three genomes of *Entamoeba* species which were currently present in GenBank, although the copy numbers are different amongst the three. PPi-PEPCK gene was also found in the genome data of a free-living amoebozoan *Mastigamoeba balamuthi* while other amoebozoans *Acanthamoeba* and *Dictyostelium* lacked the gene. In addition to the two amoebozoan genera bearing PPi-PEPCK, PPi-PEPCK homologs were also present in some eukaryotes such as excavates, dinoflagellates, diatoms, and cryptophytes, which are distantly related to each other (39). In bacteria, limited number of species in Actinobacteria including *P. freudenreichii* used in this study, Verrucomicrobia, Bacteroidetes, Planctomycetes, Proteobacteria, Firmicutes and Spirochates possessed PPi-PEPCK homologs (Table 5): Many species were found to lack PPi-PEPCK even if many of them are closely related to PPi-PEPCK-bearing species.

DISCUSSIONS

In the present study, we successfully identified the gene encoding PPi-PEPCK, which is one of the key enzymes in the central metabolism and connects sugar and organic acids metabolisms, from *E. histolytica* and *P. freudenreichii*. We also found EhPEPCKs' homologs were distributed in limited but diverse lineages of unicellular eukaryotes and bacteria, but not in archaea.

In *E. histolytica* which lacks ATP/GTP-PEPCK and PEP carboxylase, PPi-PEPCK has been believed to work for the PPi- and oxaloacetate-producing direction (27,40). This is because supply of PPi is required to replenish PPi consumed by the PPi-utilizing enzymes in glycolysis, PPi-PFK and PPDK (4,12,41). However, it is uncertain whether PPi produced by PPi-PEPCK is truly essential because PPi is supplied as a byproduct of many reactions *in vivo* (3,42). Alternatively, it would be possible that PPi-PEPCK works for the PPi-utilizing direction when there is enough supply of PPi because of the following reason. *E. histolytica* is able to consume asparagine and aspartate from media (43), and these amino acids can be converted to oxaloacetate by aminotransferase (Fig. 5) (44). Although the oxaloacetate catabolic pathway has not been experimentally confirmed in this organism which lacks TCA cycle, the ameba may be capable of producing ATP from these amino acids by using the energy of PPi if oxaloacetate is further converted to PEP by PPi-PEPCK and then pyruvate by PPDK. Future experiments to investigate which direction is critical (or whether both directions are important) for viability in *E. histolytica* would give us deeper insight into the central metabolic pathway in this human parasite.

E. histolytica has three isoforms of PPi-PEPCK which are strongly suggested to form hetero-trimers among the isoforms. This may be worth to mention because the two other species of *Entamoeba* have two isoforms and the four strains of the other eukaryotic human gut parasite *Giardia intestinalis*, for example, possess the single copy of this gene in the genomes (Fig. 6). Thus, it is still a key question whether each isoform of PPi-PEPCK from *E. histolytica* has different roles in the hetero-trimer and whether their roles are, if different, also shared amongst *Entamoeba* species. These issues should be examined in future to consider the roles of PPi-PEPCK in the human parasite *E. histolytica*.

Most of the organisms possessing PPi-PEPCK homologs have canonical ATP/GTP-PEPCK genes and some of them have PEP carboxylases as well (Table 4), although *E. histolytica* possesses the three cytosolic PPi-PEPCKs but lacks the others. How these functional homologs are utilized/regulated in each organism is an open question: They might work for different directions to each other, or might

work in different environmental conditions. Especially, in some eukaryotes, these functional homologs might be utilized in different cellular compartments. For instance, a potential PPi-PEPCK from the unicellular alga *Thalassiosira pseudonana* (XP_002285950.1) is predicted to be cytosolic according to the *in silico* predictions (data not shown) as well as those of *E. histolytica*, while its ATP-PEPCK was confirmed to exist in mitochondria and two types of PEP carboxylases localize in mitochondria and the chloroplast, respectively (45). Although the localization of ATP-PEPCK and PEP carboxylases indicates PEP–oxaloacetate conversion occurs only in organelles of the diatom, our findings suggest that this reaction important for the central carbon metabolism can also occur in the cytosol of *T. pseudonana*. Our findings in this study would further open a new era for deeper understanding the central metabolism of various organisms.

Identification of PPi-PEPCK also allows us to compare the amino acid sequences between PPi-PEPCK and the functional homologs, ATP-/GTP-PEPCK and PEP carboxylase. PPi-PEPCK bears none of the known catalytic domains such as oxaloacetate binding sites, kinase 1a/P-loop, kinase-2 region, and nucleotide binding sites conserved in ATP- and GTP-PEPCK (22). In addition, no similarity in primary structure was observed between EhPEPCK and the functional homologs. Therefore, it is unlikely that PPi-PEPCK shares the common origin with ATP- and GTP-PEPCK/ PEP carboxylase. This is similar to the case between pyruvate kinase and PPDK/PEP synthase, but different from the cases of ATP-PFK/PPi-PFK and PPDK/PEP synthase that have most probably emerged from the single origins. Our findings further illuminate the complexity of evolution amongst the proteins, which are functionally redundant but utilize different phosphate donors (e.g., PPi and nucleotide triphosphates).

Origins and evolutionary histories of PPi-PEPCKs were one of the most interesting points for authors. However, it was highly difficult to reconstruct a rooted tree of PPi-PEPCK because we could not find any proteins with sister relationships to PPi-PEPCK. In addition, whether the current distribution of PPi-PEPCK in the tree of life has shaped by lateral gene transfer or vertical inheritance followed by independent gene loss basically remains unclear because of the low resolution of the PPi-PEPCK tree (Fig. 6) with one exception for PPi-PEPCK in *Dehalobacter* sp.; the *Dehalobacter* homologue was nested in the clade of Actinobacteria although *Dehalobacter* belongs to Firmicutes, strongly suggesting actinobacterium-to-*Dehalobacter* lateral gene transfer.

In summary, PPi-PEPCK was identified at molecular level for the first time since the activity was first demonstrated more than 50 years ago (24). It was also demonstrated that PPi-PEPCK was distributed in phylogenetically diverse but limited unicellular eukaryotes and bacteria, which would remodel the central metabolic pathways in various organisms. PPi-PEPCK has no clear evolutionary relationship with ATP- or GTP-PEPCK, in sharp contrast to the case of PPi- and ATP-PFK, and that of PPDK and PEP synthase, which have evolved from the single origins. These data suggest that, in order to substitute the energy and phosphate donors from PPi to nucleotide triphosphates, proteins have employed various strategies.

Acknowledgments– We thank Dr. Nicholas E. Sherman, W.M. Keck Biomedical Mass Spectrometry Laboratory, University of Virginia, USA for mass spectrometric analysis.
This work was supported in part by Grants-in-Aid for Scientific Research on Innovative Areas (Nos. 23117001 and 23117005), Grants-in-Aid for Scientific Research (B) (No. 26293093) from Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Research Program on Emerging and Re-emerging Infectious Diseases from Japan Agency for Medical Research and Development (AMED), Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan International Cooperation Agency (JICA) and AMED, to T.N., Grant-in-Aid for Young Scientists (B) (No. 26860275) from the Japan Society for the Promotion of Science (JSPS), Japan, to Y.C., and Grant-in-Aid

for Scientific Research on Innovative Areas (No. 23117006), for Young Scientists (A) (No. 15H05606) and for Challenging Exploratory Research (No. 15K14591) from JSPS to R.K. R.K. was also supported by the Institute for Fermentation, Osaka, Japan.

Conflict of interest– The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions– Y.C. and T.N. designed this study. Y.C. performed biochemical analysis. K.N.T., and Y.S.N. designed molecular experiment. R.K. and Y.C. performed the phylogenetic analysis. Y.C., R.K. and T.N. wrote the manuscript. All the authors reviewed the results and approved the final version of the manuscript.

References

1. Liu, C. L., Hart, N., and Peck, H. D. (1982) Inorganic pyrophosphate: energy source for sulfate–reducing bacteria of the genus *Desulfotomaculum*. *Science***217**, 363–364
2. Gest, H. (1972) Energy conversion and generation of reducing power in bacterial photosynthesis. *Adv. Microb. Physiol.***7**, 243–282
3. Baykov, A. A., Malinen, A. M., Luoto, H. H., and Lahti, R. (2013) Pyrophosphate–fueled Na^+ and H^+ transport in prokaryotes. *Microbiol. Mol. Biol. Rev.***77**, 267–276
4. Mertens, E. (1991) Pyrophosphate–dependent phosphofructokinase, an anaerobic glycolytic enzyme? *FEBS lett.***285**, 1–5
5. Alves, A., Meijer, W. G., Vrijbloed, J. W., and Dijkhuizen, L. (1996) Characterization and phylogeny of the *pfp* gene of *Amycolatopsis methanolica* encoding PPi–dependent phosphofructokinase. *J. Bacteriol.***178**, 149–155
6. Siebers, B., Klenk, H. P., and Hensel, R. (1998) PPi–Dependent Phosphofructokinase from *Thermoproteus tenax*, an Archaeal Descendant of an Ancient Line in Phosphofructokinase Evolution. *J. Bacteriol.***180**, 2137–2143
7. Mertens, E., Ladror, U. S., Lee, J. A., Miretsky, A., Morris, A., Rozario, C., Kemp, R. G., and Müller, M. (1998) The pyrophosphate–dependent phosphofructokinase of the protist, *Trichomonas vaginalis*, and the evolutionary relationships of protist phosphofructokinases. *J. Mol. Evol.***47**, 739–750
8. Chi, A., and Kemp, R. G. (2000) The primordial high energy compound: ATP or inorganic pyrophosphate? *J. Biol. Chem.* **275**, 35677–35679
9. Müller, M., Lee, J. A., Gordon, P., Gaasterland, T., and Sensen, C. W. (2001) Presence of prokaryotic and eukaryotic species in all subgroups of the PPi–dependent group II phosphofructokinase protein family. *J. Bacteriol.***183**, 6714–6716
10. Bapteste, E., Moreira, D., and Philippe, H. (2003) Rampant horizontal gene transfer and phospho–donor change in the evolution of the phosphofructokinase. *Gene***318**, 185–191
11. Pocalyko, D. J., Carroll, L. J., Martin, B. M., Babbitt, P. C., and Dunaway–Mariano, D. (1990) Analysis of sequence homologies in plant and bacterial pyruvate phosphate dikinase, enzyme I of the bacterial phospoenolpyruvate: sugar phosphotransferase system and other PEP–utilizing enzymes. Identification of potential catalytic and regulatory motifs. *Biochemistry***29**, 10757–10765
12. Reeves, R. E., South, D. J., Blytt, H. J., and Warren, L. G. (1974) Pyrophosphate: D–Fructose 6–Phosphate 1–Phosphotransferase A NEW ENZYME WITH THE GLYCOLYTIC FUNCTION OF 6–PHOSPHOFRUCTOKINASE. *J. Biol. Chem.***249**, 7737–7741
13. O'Brien, W. E., Bowien, S., and Wood, H. G. (1975) Isolation and characterization of a

pyrophosphate–dependent phosphofructokinase from *Propionibacterium shermanii*. *J. Biol. Chem.***250**, 8690–8695

14. van der Merwe, M. J., Groenewald, J.–H., Stitt, M., Kossmann, J., and Botha, F. C. (2010) Downregulation of pyrophosphate: D–fructose–6–phosphate 1–phosphotransferase activity in sugarcane culms enhances sucrose accumulation due to elevated hexose–phosphate levels. *Planta* **231**, 595–608
15. Michels, P. A., Chevalier, N., Opperdoes, F. R., Rider, M. H., and Rigden, D. J. (1997) The Glycosomal ATP‐Dependent Phosphofructokinase of *Trypanosoma brucei* must have Evolved from an Ancestral Pyrophosphate‐Dependent Enzyme. *Eur. J. Biochem.***250**, 698–704
16. Niersbach, M., Kreuzaler, F., Geerse, R., Postma, P., and Hirsch, H. (1992) Cloning and nucleotide sequence of the *Escherichia coli* K–12 *ppsA* gene, encoding PEP synthase. *Mol. Gen. Genet.***231**, 332–336
17. Davies, D. (1979) The central role of phosphoenolpyruvate in plant metabolism. *Annu. Rev. Plant. Physiol.***30**, 131–158
18. Sauer, U., and Eikmanns, B. J. (2005) The PEP–pyruvate–oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol. Rev.***29**, 765–794
19. Siu, P. M., and Wood, H. G. (1962) Phosphoenolpyruvic carboxytransphosphorylase, a CO_2 fixation enzyme from propionic acid bacteria. *J. Biol. Chem.***237**, 3044–3051
20. Bandurski, R. S., and Greiner, C. M. (1953) The enzymatic synthesis of oxaloacetate from phosphoryl–enolpyruvate and dioxide. *J. Biol. Chem.***204**, 781-786
21. Fukuda, W., Fukui, T., Atomi, H., and Imanaka, T. (2004) First characterization of an archaeal GTP–dependent phosphoenolpyruvate carboxykinase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.***186**, 4620–4627
22. Aich, S., and Delbaere, L. T. (2007) Phylogenetic study of the evolution of PEP–carboxykinase. *Evol. Bioinform Online.***3**, 333-340
23. Matte, A., Tari, L. W., Goldie, H., and Delbaere, L. T. (1997) Structure and mechanism of phosphoenolpyruvate carboxykinase. *J. Biol. Chem.***272**, 8105–8108
24. Siu, P. M., Wood, H. G., and Stjernholm, R. L. (1961) Fixation of CO_2 by phosphoenolpyruvic carboxytransphosphorylase. *J. Biol. Chem.***236**, PC21–PC22
25. Lochmüller, H., Wood, H. G., and Davis, J. J. (1966) Phosphoenolpyruvate Carboxytransphosphorylase II. CRYSTALLIZATION AND PROPERTIES. *J. Biol. Chem.***241**, 5678–5691
26. Haberland, M. E., Willard, J. M., and Wood, H. G. (1972) Phosphoenolpyruvate carboxytransphosphorylase. Study of the catalytic and physical structures. *Biochemistry***11**, 712–722
27. Reeves, R. E. (1970) Phosphopyruvate carboxylase from *Entamoeba histolytica*. *Biochim. Biophys. Acta.***220**, 346–349
28. Diamond, L. S., Harlow, D. R., and Cunnick, C. C. (1978) A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Trans. R. Soc. Trop. Med. Hyg.***72**, 431–432
29. Marumo, K., Nakada–Tsukui, K., Tomii, K., and Nozaki, T. (2014) Ligand heterogeneity of the cysteine protease binding protein family in the parasitic protist *Entamoeba histolytica*. *Int. J. Parasitol.* **44**, 625-35
30. Nakada‐Tsukui, K., Okada, H., Mitra, B. N., and Nozaki, T. (2009) Phosphatidylinositol‐phosphates mediate cytoskeletal reorganization during phagocytosis via a unique modular protein consisting of RhoGEF/DH and FYVE domains in the parasitic protozoon *Entamoeba histolytica*. *Cell. Microbiol.***11**, 1471–1491

31. Nozaki, T., Asai, T., Sanchez, L. B., Kobayashi, S., Nakazawa, M., and Takeuchi, T. (1999) Characterization of the Gene Encoding Serine Acetyltransferase, a Regulated Enzyme of Cysteine Biosynthesis from the Protist Parasites *Entamoeba histolytica* and *Entamoeba dispar* REGULATION AND POSSIBLE FUNCTION OF THE CYSTEINE BIOSYNTHETIC PATHWAY IN ENTAMOEBA. *J. Biol. Chem.***274**, 32445–32452
32. Nozaki, T., Asai, T., Kobayashi, S., Ikegami, F., Noji, M., Saito, K., and Takeuchi, T. (1998) Molecular cloning and characterization of the genes encoding two isoforms of cysteine synthase in the enteric protozoan parasite *Entamoeba histolytica*. *Mol. Biochem. Parasitol.***97**, 33–44
33. Mi–Ichi, F., Yousuf, M. A., Nakada–Tsukui, K., and Nozaki, T. (2009) Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc. Natl. Acad. Sci. U S A* **106**, 21731–21736
34. Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral–Zettler, L. A., Armbrust, E. V., Archibald, J. M., Bharti, A. K., and Bell, C. J. (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.***12**, e1001889
35. Katoh, K., Misawa, K., Kuma, K. i., and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.***30**, 3059–3066
36. Stamatakis, A. (2006) RAxML–VI–HPC: maximum likelihood–based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics***22**, 2688–2690
37. Medina, V., Pontarollo, R., Glaeske, D., Tabel, H., and Goldie, H. (1990) Sequence of the *pckA* gene of *Escherichia coli* K–12: relevance to genetic and allosteric regulation and homology of *E. coli* phosphoenolpyruvate carboxykinase with the enzymes from *Trypanosoma brucei* and *Saccharomyces cerevisiae*. *J. Bacteriol.***172**, 7151–7156
38. Croniger, C. M., Olswang, Y., Reshef, L., Kalhan, S. C., Tilghman, S. M., and Hanson, R. W. (2002) Phosphoenolpyruvate carboxykinase revisited: Insights into its metabolic role. *Biochem. Mol. Biol. Educ.***30**, 14–20
39. Adl, S. M., Simpson, A. G., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., Brown, M. W., Burki, F., Dunthorn, M., and Hampl, V. (2012) The revised classification of eukaryotes. *J. Eukaryot. Microbiol.***59**, 429–514
40. Müller, M., Mentel, M., van Hellemond, J. J., Henze, K., Woehle, C., Gould, S. B., Yu, R.–Y., van der Giezen, M., Tielens, A. G., and Martin, W. F. (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.***76**, 444–495
41. Reeves, R. E. (1968) A new enzyme with the glycolytic function of pyruvate kinase. *J. Biol. Chem.***243**, 3202–3204
42. Kulaev, I. S., and Vagabov, V. M. (1983) Polyphosphate metabolism in microorganisms. *Adv. Microb. Physiol.***24**, 83–171
43. Zuo, X., and Coombs, G. H. (1995) Amino acid consumption by the parasitic, amoeboid protists *Entamoeba histolytica* and *E. invadens*. *FEMS Microbiol. Lett.***130**, 253–258
44. Clark, C. G., Alsmark, U. C., Tazreiter, M., Saito–Nakano, Y., Ali, V., Marion, S., Weber, C., Mukherjee, C., Bruchhaus, I., Tannich, E., Leippe, M., Sicheritz–Ponten, T., Foster, P. G., Samuelson, J., Noel, C. J., Hirt, R. P., Embley, T. M., Gilchrist, C. A., Mann, B. J., Singh, U., Ackers, J. P., Bhattacharya, S., Bhattacharya, A., Lohia, A., Guillen, N., Duchene, M., Nozaki, T., and Hall, N. (2007) Structure and content of the *Entamoeba histolytica* genome. *Adv. Parasitol.***65**, 51–190
45. Tanaka, R., Kikutani, S., Mahardika, A., and Matsuda, Y. (2014) Localization of enzymes relating to C4 organic acid metabolisms in the marine diatom, *Thalassiosira pseudonana*. *Photosynth. Res.* **121**, 251–63
46. Pineda, E., Encalada, R., Vázquez, C., González, Z., Moreno–Sánchez, R., and Saavedra, E. (2015) Glucose Metabolism and Its Controlling Mechanisms in *Entamoeba histolytica.* in

Amebiasis, Springer. pp 351–372
47. Jeelani, G., and Nozaki, T. (2014) Metabolomic analysis of *Entamoeba*: Applications and implications. *Curr. Opin. Microbiol* **20**, 118–124

Figure legends

Figure 1. SDS-PAGE analysis of purified native EhPEPCKs (A) and *E. coli*-expressed recombinant EhPEPCK1, EhPEPCK2, and PfPEPCK (B).
A, Relative PPi-PEPCK activities in the selected MonoQ fractions (per volume) are shown on the top panel. Approximately 15 µl of the MonoQ fractions were applied to SDS-PAGE and stained with Coomassie Brilliant Blue (bottom). The arrow indicates the band corresponding to EhPEPCK. B, Homogeneity of purified *E. coli*-expressed recombinant EhPEPCK1 (labeled as PEPCK1), EhPEPCK2 (PEPCK2), and PfPEPCK. One µg of recombinant proteins after the purification by MonoQ column was subjected to SDS-PAGE and Coomassie Brilliant Blue staining.

Figure 2. Molecular weight of purified native EhPEPCKs.
A, Chromatograph of native EhPEPCKs on Superdex200. The optical absorbance at 280 nm is shown with a thick line and the relative activity per volume of each fraction is shown with squares and a thin line. The highest activity per volume in all fractions was defined as 100%. B, Estimation of molecular weight of the proteins corresponding to the two peaks in "A" (a circle and a triangle). Standard proteins are shown with diamonds.

Figure 3. WB analysis of EhPEPCKs with anti-HA (A) and anti-EhPEPCK (B and C) antibodies.
A, Expression of EhPEPCK1-HA and 2-HA was confirmed. Note that proteins with the expected size (~130 kDa) were expressed. Amebic lysates containing 5 µg of protein were applied per lane on SDS-PAGE. B, Expression of EhPEPCKs in EhPEPCK1 or 2-HA expressing and mock control *E. histolytica* transformants. Total lysates of the above mentioned transformants (0.2-5 µg/ lane) and purified histidine-tagged PEPCK1 and 2 (2-50 ng/ lane) were electrophoresed on SDS-PAGE and subjected to WB analysis with anti-EhPEPCK antiserum. Note that lysates from mock transformant showed two bands while EhPEPCK1-HA expressing transformant showed one additional band (shown with an asterisk) corresponding to PEPCK1 with the HA-tag. Recombinant PEPCKs possesses the His-tag at the amino terminus and were purified using Ni-NTE and MonoQ columns to homogeneity. C, Specificity of anti-EhPEPCK antibody. Approximately 1 µg of lysate from the wild-type *E. histolytica* was reacted with anti-EhPEPCK antibody.

Figure 4. Distribution of *E. histolytica* PEPCKs by fractionation and WB analysis.
Trophozoites of mock, EhPEPCK1 or 2-HA-expressing transformants were mechanically disrupted, and centrifuged to separate into cytosolic and organelle fractions. Those fractions were reacted with anti-EhPEPCK, cysteine synthase 1 (CS1, cytosolic marker), or Cpn60 (a mitosomal matrix protein, used as an organellar marker) antibodies. Whereas approximately 5 µg of lysates was used for anti-PEPCK antibody, 20 µg of lysates were reacted with anti-CS1 and anti-Cpn60 antibodies.

Figure 5. Glycolytic pathway and oxaloacetate—malate node of *E. histolytica*.
Arrows indicate the reaction steps and their directions conducted by the enzymes in *E. histolytica* (46,47). PFK: phosphofructo kinase, PPDK: pyruvate phosphate dikinase, PEPCK: PEP carboxykinase.

Figure 6. Phylogenic tree of PPi-PEPCKs and their homologs. Sequences were aligned by MAFFT and the dataset was subjected to the maximum likelihood analysis with RAxML ver. 7.2.8. Only bootstrap supports equal to or more than 85 are shown on each node. White letters in black boxes and black letters indicate eukaryotic and bacterial phyla, respectively.

Table 1
Purification of PPi-PEPCK from *E. histolytica*

	Activity (U)	Specific activity (U mg^{-1})	Puridication (-fold)	Yield (%)
100k×*g* Sup.	75.4	0.165	1.0	100
Butyl toyopeal	32.1	4.86	29.4	42.6
MonoQ	8.06	34.0	206	10.7

1U= Oxidation of 1 μ mol NADH min^{-1}

Table 2

Proteins detected by MS/MS analysis

Identified Proteins		Accession Number		Molecular Weight (kDa)	Quantitative value
hypothetical protein	(EhPEPCK1)	EHI_166920	XP_654765	131	214
hypothetical protein	(EhPEPCK2)	EHI_030750	XP_650862	131	209
hypothetical protein	(EhPEPCK3)	EHI_198620	XP_655201	131	198
hypothetical protein		EHI_192430	XP_656829	29	1
hypothetical protein		EHI_189920	XP_652326	62	1

Peptide Thresholds: Standard

Protein Thresholds: 90.0% minimum and 1 peptide minimum

Table 3

Relative quantification of EhPEPCKs isotypes

	Peak area of monoisotopic ion			Normalized relative ratio against common peptides		
Peptide	EhPEPCK1 -HA IP	EhPEPCK2 -HA IP	Purification from cells	EhPEPCK1 -HA IP	EhPEPCK2 -HA IP	Purification from cells
Unique to EhPEPCK1						
GTDYPILNIQELEALADLK	4.54E+07	3.96E+06	4.61E+06			
SDAIDVVAPLVDIIAEGDQESTAP IDAR	4.09E+07	6.45E+06	2.17E+06			
NISDAIFEGK	2.41E+09	3.91E+08	6.71E+08			
Sum	2.50E+09	4.01E+08	6.78E+08	3.0E-01	5.5E-02	1.1E-01
Specific to EhPEPCK2 and 3						
ITTAFANHFLR	9.22E+07	6.07E+08	1.68E+08			
KLESLANLK	3.81E+08	1.76E+09	8.54E+08			
LAGPLLEEVEESEVNHTTAPIDAR	2.62E+08	1.03E+09	2.54E+08			
Sum	7.35E+08	3.40E+09	1.28E+09	8.8E-02	4.6E-01	2.1E-01
Unique to EhPEPCK2						
AVQEIFDHDFSKR	5.79E+06	2.38E+07	2.05E+07	6.9E-04	3.3E-03	3.3E-03
Unique to EhPEPCK3						
GANLSSQYLR	2.14E+08	9.04E+06	3.37E+08	2.6E-02	1.2E-03	5.5E-02
Common to all						
YLIEHGYLEPCPDVTYNGK	2.15E+08	1.87E+08	3.52E+08			
LFQRPDDAVFR	4.41E+09	4.39E+09	2.91E+09			
DFFPAAK	3.71E+09	2.74E+09	2.92E+09			
Sum	8.34E+09	7.32E+09	6.18E+09	1.0	1.0	1.0

Table 4
Distribution of PPi-PEPCK and functional homologues in 10 represented genomes

Organism	Phylum	Identity (%)*	PEPCK ATP/GTP	PEP carboxylase
Eukaryote				
Entamoeba histolytica HM-1:IMSS	Amoebozoa	100	-	-
Naegleria gruberi strain ATCC 30224	Heterolobosea	45	ATP	-
Guillardia theta CCMP2712	Cryptophyta	44	ATP	+
Thalassiosira pseudonana CCMP1335	Heterokontophyta	44	ATP	+
Giardia intestinalis strain ATCC 50581	Metamonada	39	GTP	-
Bacteria				
Pedosphaera parvula Ellin514	Verrucomicrobia	48	GTP	-
Planctomyces maris DSM 8797	Planctomycetes	46	ATP	-
Thioalkalivibrio nitratireducens DSM 14787	Proteobacteria	46	GTP	+
Actinomyces graevenitzii C83	Actinobacteria	45	GTP**	-
Formosa agariphila KMM 3901	Bacteroidetes	42	ATP	+**

One organism from every phylum which has a PPi-PEPCK homolog with the highest identity to EhPEPCK in blastp search was listed. For bacteria, only phylum in which more than two organisms possess PPi-PEPCK homolog were shown.

+, ATP, and GTP in the table indicates that the organism has a protein which shows 35% or higher identity to biochemically analyzed PEP carboxylase (from *E. coli*, UniProtKB; P00864), ATP-PEPCK (from *E. coli*, UniProtKB; P22259), and GTP-PEPCK (from *Giardia intestinalis*, XP_001709869.1), respectively.

* Identity against EhPEPCK1

** A protein with significant homology to the query while the identity is lower than 35%

Table 5
Distribution of PPi-PEPCK in Bacteria

Phylum	No. of genomes[*1]	No. of PPi-PEPCK containing genome*1
Actinobacteria	910	104
Aquificae	16	0
Armatimonadetes	3	0
Bacteroidetes/Chlorobi group	470	19
Caldiserica	2	0
Chlamydiae/Verrucomicrobia group	59	15
Chloroflexi	25	0
Chrysiogenetes	2	0
Cyanobacteria	104	0
Deferribacteres	6	0
Deinococcus-Thermus	43	0
Dictyoglomi	2	0
Elusimicrobia	2	0
Fibrobacteres/Acidobacteria group	46	1
Firmicutes	1129	1
Fusobacteria	25	0
Gemmatimonadetes	5	0
Nitrospinae	2	0
Nitrospirae	10	0
Planctomycetes	22	18
Proteobacteria	2200	33
Spirochaetes	80	2
Synergistetes	18	0
Tenericutes	140	0
Thermodesulfobacteria	7	0
Thermotogae	24	0
Unclassified Bacterium	210	1

*1: at 5th May 2015 in GenBank.

Fig. 1

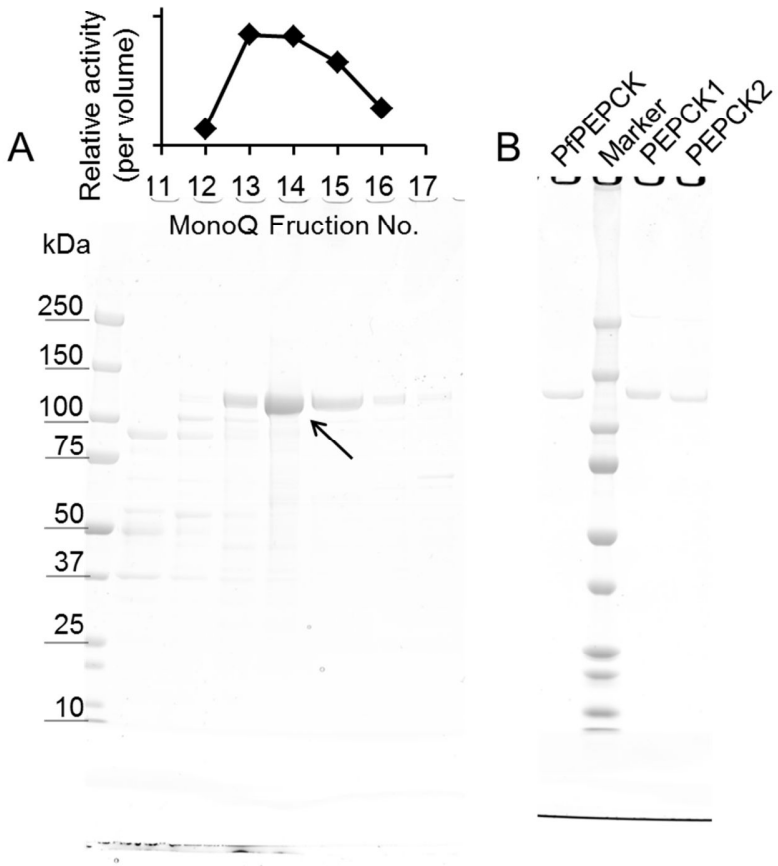

A
Relative activity
(per volume)
11 12 13 14 15 16 17
MonoQ Fruction No.
kDa
250
150
100
75
50
37
25
10
B
PfPEPCK
Marker
PEPCK1
PEPCK2

Fig. 2

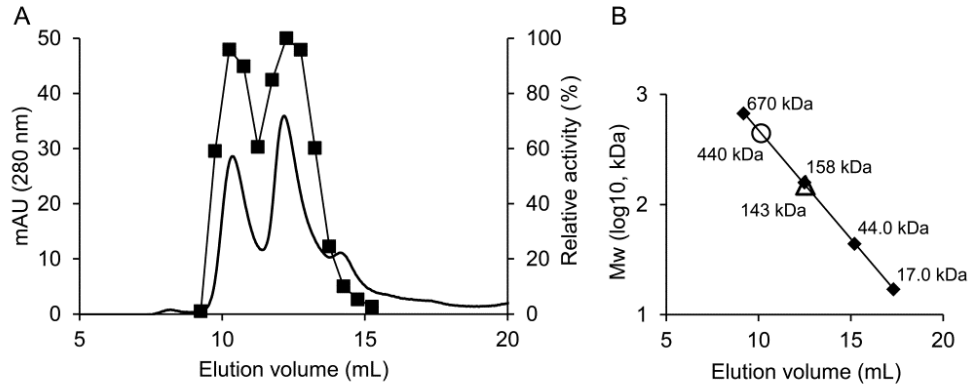

A
mAU (280 nm)
Relative activity (%)
Elution volume (mL)
B
Mw (log10, kDa)
670 kDa
440 kDa
158 kDa
143 kDa
44.0 kDa
17.0 kDa
Elution volume (mL)

Fig. 3

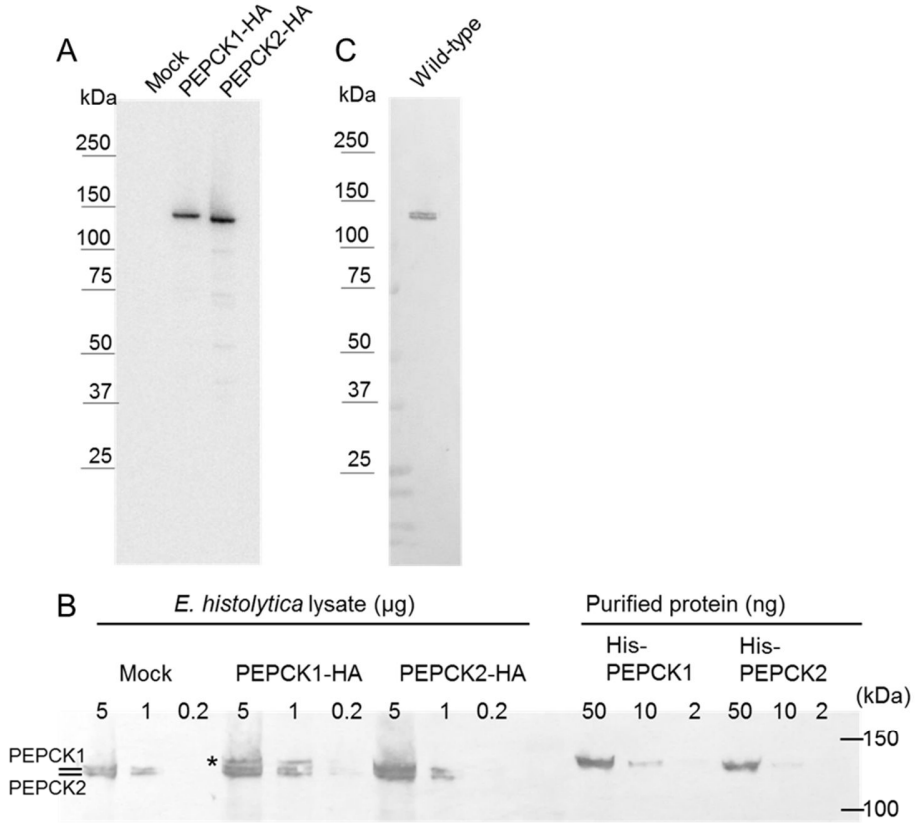

A
Mock
PEPCK1-HA
PEPCK2-HA
kDa
250
150
100
75
50
37
25
C
Wild-type
kDa
250
150
100
75
50
37
25
B
E. histolytica lysate (μg)
Purified protein (ng)
Mock
PEPCK1-HA
PEPCK2-HA
His-PEPCK1
His-PEPCK2
5 1 0.2 5 1 0.2 5 1 0.2
50 10 2 50 10 2
(kDa)
150
100
PEPCK1
PEPCK2
*

Fig. 4

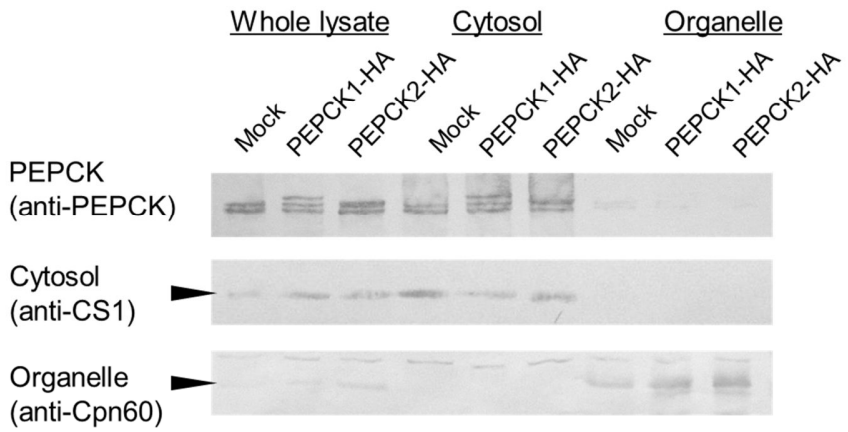

Whole lysate
Cytosol
Organelle
Mock
PEPCK1-HA
PEPCK2-HA
Mock
PEPCK1-HA
PEPCK2-HA
Mock
PEPCK1-HA
PEPCK2-HA
PEPCK
(anti-PEPCK)
Cytosol
(anti-CS1)
Organelle
(anti-Cpn60)

Fig. 5

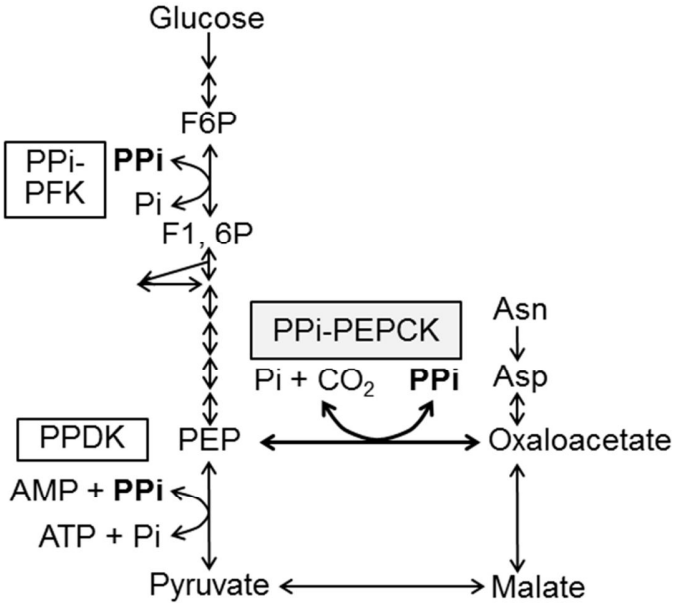
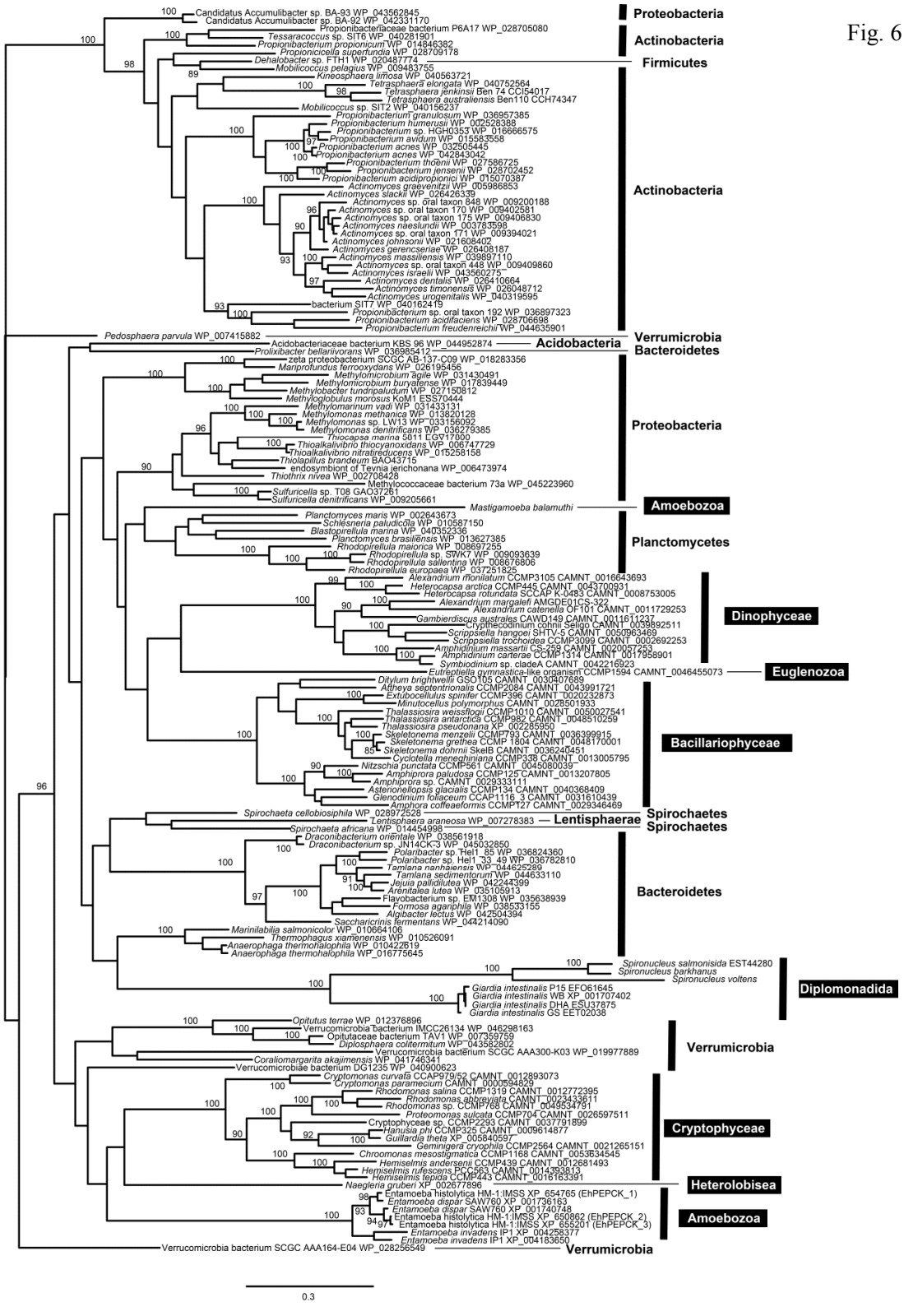

Glucose
F6P
PPi-
PFK
PPi
Pi
F1, 6P
PPi-PEPCK
Pi + CO_2
PPi
Asn
Asp
PPDK
PEP
Oxaloacetate
AMP + PPi
ATP + Pi
Pyruvate
Malate

Fig. 6


Proteobacteria
Actinobacteria
Firmicutes
Actinobacteria
Verrumicrobia
Acidobacteria
Bacteroidetes
Proteobacteria
Amoebozoa
Planctomycetes
Dinophyceae
Euglenozoa
Bacillariophyceae
Spirochaetes
Lentisphaerae
Spirochaetes
Bacteroidetes
Diplomonadida
Verrumicrobia
Cryptophyceae
Heterolobisea
Amoebozoa
Verrumicrobia
0.3